4.

(i) $L_{\pi_{\theta_1}}(\pi_\theta) = \eta(\pi_{\theta_1}) + \sum_{s\in S} d_\mu^{\pi_{\theta_1}}(s) \sum_{a\in A} \pi_{\theta_1}(a|s) A^{\pi_{\theta_1}}(s,a)$

$\sum_{a\in A} \pi_{\theta_1}(a|s) A^{\pi_{\theta_1}}(s,a) = \sum_{a\in A} \pi_{\theta_1}(a|s) Q^{\pi_{\theta_1}}(s,a) - \sum_{a\in A} \pi_{\theta_1}(a|s) V^{\pi_{\theta_1}}(s)$

$= V^{\pi_{\theta_1}}(s) - V^{\pi_{\theta_1}}(s) = 0$

$\Rightarrow L_{\pi_{\theta_1}}(\pi_{\theta_1}) = \eta(\pi_{\theta_1})$ #

(ii) $\nabla_\theta L_{\pi_{\theta_1}}(\pi_\theta) = \nabla_\theta \eta(\pi_{\theta_1}) + \nabla_\theta \sum_{s\in S} d_\mu^{\pi_{\theta_1}}(s) \sum_{a\in A} \pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)$

($\nabla_\theta \eta(\pi_{\theta_1}) = 0$, $\nabla_\theta d_\mu^{\pi_{\theta_1}}(s) = 0$, $\nabla_\theta A^{\pi_{\theta_1}}(s,a) = 0$)

$= \sum_{s\in S} d_\mu^{\pi_{\theta_1}}(s) \sum_{a\in A} A^{\pi_{\theta_1}}(s,a) \nabla_\theta \pi_\theta(a|s)$

$\nabla_\theta \eta(\pi_\theta) = \nabla_\theta \eta(\pi_{\theta_1}) + \sum_{s\in S} \nabla_\theta d_\mu^{\pi_\theta}(s) \sum_{a\in A} \pi_\theta(a|s) A^{\pi_{\theta_1}}(s,a)$

($\nabla_\theta \eta(\pi_{\theta_1}) = 0$)

$+ \sum_{s\in S} d_\mu^{\pi_\theta}(s) \sum_{a\in A} A^{\pi_{\theta_1}}(s,a) \nabla_\theta \pi_\theta(a|s)$

$\nabla_\theta \eta(\pi_\theta)\big|_{\theta=\theta_1} = \sum_{s\in S} \nabla_\theta d_\mu^{\pi_\theta}(s)\big|_{\theta=\theta_1} \sum_{a\in A} \pi_{\theta_1}(a|s) A^{\pi_{\theta_1}}(s,a)$

($\sum_{a\in A} \pi_{\theta_1}(a|s) A^{\pi_{\theta_1}}(s,a) = 0$)

$+ \sum_{s\in S} d_\mu^{\pi_\theta}(s) \sum_{a\in A} A^{\pi_{\theta_1}}(s,a) \nabla_\theta \pi_\theta(a|s)\big|_{\theta=\theta_1}$

$= \sum_{s\in S} d_\mu^{\pi_{\theta_1}}(s) \sum_{a\in A} A^{\pi_{\theta_1}}(s,a) \nabla_\theta \pi_\theta(a|s)\big|_{\theta=\theta_1}$

$= \nabla_\theta L_{\pi_{\theta_1}}(\pi_\theta)\big|_{\theta=\theta_1}$ #

2.

(a) $D(\lambda) = \min_{\theta \in \mathbb{R}^d} -(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T(\theta-\theta_k) + \lambda(\frac{1}{2}(\theta-\theta_k)^T H(\theta-\theta_k) - \delta)$

$= \min_{\theta \in \mathbb{R}^d} (\frac{\lambda}{2} H(\theta-\theta_k) - \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T(\theta-\theta_k) - \lambda\delta$

at minimum,

$\nabla_\theta L(\theta,\lambda) = \nabla_\theta (\frac{\lambda}{2} H(\theta-\theta_k) - \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T(\theta-\theta_k) = 0$

$= \frac{\lambda}{2} H(\theta-\theta_k) + \frac{\lambda}{2} H(\theta-\theta_k) - \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k} = 0$

$(\theta-\theta_k) = \frac{1}{\lambda} H^{-1} \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}$

$D(\lambda) = (\frac{1}{2}\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k} - \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T \frac{1}{\lambda} H^{-1} \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k} - \lambda\delta$

$= -\frac{1}{2\lambda}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T H^{-1}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}) - \lambda\delta$ **

$\nabla_\lambda D(\lambda) = \frac{1}{2\lambda^2}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T H^{-1}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k}) - \delta = 0$

$2\lambda^2\delta = (\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T H^{-1}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})$

$\lambda^* = \sqrt{\frac{1}{2\delta}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T H^{-1}(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})}$ *

(b)

$L(\theta,\lambda^*) = -(\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k})^T(\theta-\theta_k) + \lambda^*(\frac{1}{2}(\theta-\theta_k)^T H(\theta-\theta_k) - \delta)$

$\nabla_\theta L(\theta,\lambda^*) = -\nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k} + \lambda^*(H(\theta-\theta_k)) = 0$

$\theta = \frac{1}{\lambda^*} H^{-1} \nabla_\theta L_{\theta_k}(\theta)\big|_{\theta=\theta_k} + \theta_k$, $\alpha = \frac{1}{\lambda^*}$ *